

EKSTRAKSI FITUR UNTUK PENGENALAN WAJAH PADA RAS MONGOLOID MENGGUNAKAN PRINCIPAL COMPONENT ANALYSIS (PCA)

Dwiandi Susantyo

Jurusan Teknik Informatika – S1, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
Jl. Nakula 1 No. 5-11 Semarang 50131, Telp (024)3520165 Home Page : www.dinus.ac.id
E-mail : sekretariat@dinus.ac.id dwiandisusantyo@yahoo.co.id

Abstract

One of the processes for face recognition is feature extraction. The purpose of feature extraction is used to taking important features of an object face. The accuracy level for feature extraction process will determine the quality of the results that will make the rest of the processes a lot easier. PCA is used to reduce a feature dimension of object, so can be used to take important characteristic of object processed. By doing some testing using PCA as taking on characteristics of the face image and euclidean distance to find the minimum distance between data training data and data test so can produce accuration values 78.89% of 90 data test.

Key word : feature extraction, PCA, euclidean distance, training data, test data

Abstrak

Salah satu proses dari pengenalan wajah adalah ekstraksi fitur. Tujuan utama dari ekstraksi fitur digunakan untuk mengambil ciri penting dari suatu objek wajah. Tingkat akurasi proses ekstraksi fitur pada wajah akan menentukan kualitas hasil sehingga akan mempermudah dalam pemrosesan selanjutnya. PCA digunakan untuk mereduksi dimensi ciri dari suatu obyek, sehingga dapat digunakan mengambil karakteristik yang penting saja dari obyek yang diproses. Dengan melakukan beberapa pengujian menggunakan PCA sebagai pengambilan ciri pada citra wajah dan *euclidean distance* untuk mencari nilai minimum jarak antar citra data training dengan cira data test menghasilkan nilai akurasi sebesar 78.89% dari 90 data test.

Kata kunci : ekstraksi fitur, PCA, *euclidean distance*, data training, data test

1. Pendahuluan

1.1. Latar Belakang

Tujuan utama dari ekstraksi fitur digunakan untuk mengambil ciri penting dari suatu objek. Hasil ekstraksi objek ini dapat dimanfaatkan pada sistem biometrika yang digunakan untuk identifikasi personal pada penggunaan mesin absensi, akses kontrol, dan lain-lain. Beberapa masalah timbul dalam proses ekstraksi yang biasa disebut *robust*. Perubahan posisi atau sudut kemiringan citra, perubahan intensitas cahaya yang

terjadi pada saat pengambilan citra wajah serta perubahan detail seperti adanya janggut, kumis, pemakaian aksesoris, perubahan gaya, perubahan ekspresi wajah menjadi tertawa, tersenyum, muram, menangis, dan lain sebagainya mengakibatkan citra dapat direpresentasikan berbeda oleh sistem. Dalam rangka ekstraksi fitur untuk pengenalan wajah sebagai tujuan dalam penelitian ini, metode *Principal Component Analysis* (PCA) diusulkan, proses ekstraksi dilakukan dengan cara pengambilan ciri dari citra wajah

dengan mereduksi dimensi ciri dari suatu obyek, sehingga ukuran dari obyek akan lebih ringkas dan mampu mengambil karakteristik yang penting saja dari objek yang diolah.

1.2. Rumusan Masalah

Dari latar belakang yang telah diuraikan, permasalahan yang harus dipecahkan dalam penelitian ini adalah :

1. Perbedaan pada ekspresi wajah, pemakaian aksesoris, perubahan intensitas cahaya, perubahan posisi ataupun sudut kemiringan citra mengakibatkan citra seseorang dapat dipresentasikan berbeda, sehingga informasi yang diperoleh menjadi tidak unik.
2. Untuk melakukan proses recognition diperlukan data awal yang uni, sehingga diperlukan tahap preprocessing pada data, oleh karena itu Principal Component Analysis digunakan pada tahap ini.

1.3. Batasan Masalah

Batasan masalah pada penelitian ini adalah :

1. Pada proses ekstraksi fitur dilakukan dengan menggunakan metode *Principal Component Analysis* (PCA).
2. Metode ini akan diimplementasikan pada wajah.
3. Citra wajah yang akan diekstraksi adalah citra wajah dalam keadaan diam dan tampak dari depan.
4. Batas sudut kemiringan citra wajah maksimal 30 derajat.
5. Jarak citra wajah yang akan diambil adalah 30 cm – 50 cm.
6. Ekstraksi fitur untuk pengenalan wajah ini hanya sebatas untuk penelitian saja dan tidak digunakan sebagai aplikasi.

1.4. Tujuan Penelitian

Tujuan yang akan dicapai dalam penelitian ini adalah memberikan penyelesaian untuk permasalahan pengenalan wajah yang hasilnya diharapkan dapat :

1. Untuk mendapatkan informasi yang unik disetiap fitur wajah, maka penelitian ini dilakukan dengan pendekatan menggunakan metode *Principal Component Analysis* (PCA).
2. Agar informasi yang dihasilkan memiliki tingkat akurasi yang optimal, maka penelitian ini ditujukan untuk membangun proses pengenalan wajah menggunakan metode *Principal Component Analysis* (PCA) agar dimungkinkan hasil yang dicapai lebih akurat.

2. Tinjauan Pustaka

2.1. Pengolahan Citra (*Image Processing*)

Pengolahan citra atau *image processing* merupakan bentuk pengolahan sinyal yang masukannya berupa gambar, sedangkan keluaran dari pengolahan gambar dapat berupa gambar atau sejumlah karakteristik yang berkaitan dengan gambar.

2.2. Konversi Citra RGB menjadi Grayscale

Proses pengubahan citra yang berwarna menjadi citra grayscale melalui beberapa tahap. Pertama adalah mengambil nilai R, G dan B dari suatu citra bertipe RGB. Pada tipe citra direpresentasikan dalam 24 bit, sehingga diperlukan proses untuk mengambil masing - masing 3 kelompok 8 bit dari 24 bit tadi. Perhitungan yang digunakan untuk mengubah citra berwarna yang mempunyai nilai matriks masing-

masing R, G, dan B menjadi citra grayscale dapat dilakukan dengan mengambil rata-rata dari nilai R, G, dan B sehingga dapat dituliskan menjadi:

$$\text{Grayscale} = \frac{R+G+B}{3}$$

2.3. Principal Component Analysis (PCA)

Dalam proses ekstraksi fitur pada pengenalan wajah dihadapkan kepada banyak variabel dengan data yang berdimensi tinggi. Operasi-operasi yang dilakukan terhadap citra wajah merupakan operasi yang dilakukan terhadap matriks yang berdimensi tinggi dan hal ini akan membutuhkan biaya komputasi yang sangat besar. Oleh karena itu, dibutuhkan *Principal Component Analysis* (PCA) yang dapat mengurangi besarnya dimensi dari data yang diobservasi menjadi dimensi yang lebih kecil tanpa kehilangan informasi yang signifikan dalam menggambarkan keseluruhan data.

2.4. Jarak Euclidean

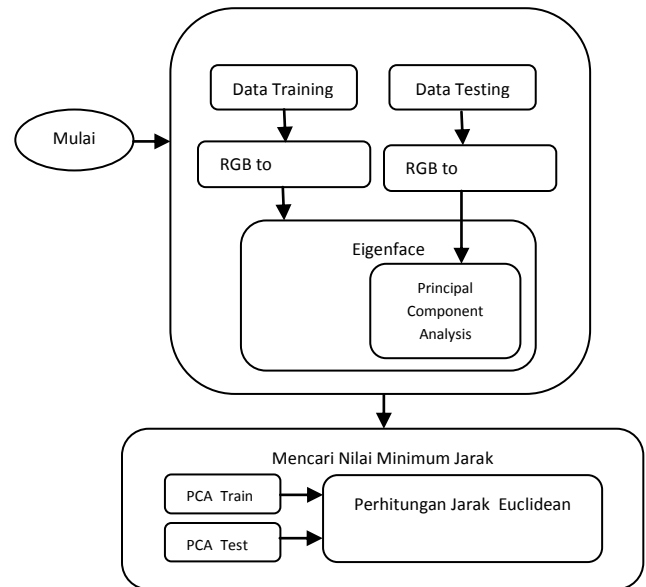
Dari hasil ekstraksi citra wajah, maka setiap citra dinyatakan dalam bentuk vektor. Untuk mengukur tingkat kemiripan antara 2 buah citra wajah dapat digunakan jarak antara kedua vektor citra wajahnya. Semakin kecil jarak 2 vektor citra wajah, semakin tinggi pula tingkat kemiripannya. Untuk mengukur jarak antara 2 vektor citra wajah tersebut, digunakanlah metode jarak Euclidean. Jarak Euclidean dapat dirumuskan sebagai berikut :

$$D = \sqrt{\sum_{i=0}^m |x_i - y_i|^2}$$

3. Metodologi Penelitian

3.1. Desain Penelitian

Desain penelitian mengacu pada tahapan proses pada sistem, yaitu input, proses dan output. Dengan input bahan penelitian berupa data training dan test, proses sebagai pemroses bahan penelitian dan output sebagai keluaran hasil dari proses.



Gambar 1. Desain Penelitian

Tahapan ekstraksi fitur pada sistem diawali dengan proses training data yaitu pembacaan data citra training yang nantinya akan menjadi database. Data training kemudian dikonversi dari citra RGB menjadi grayscale. Tujuan dari proses ini untuk mereduksi citra pada data training sehingga mudah diolah karena mengandung nilai yang lebih sedikit yaitu 8 bit daripada citra RGB dengan 24 bit warna. Tahap selanjutnya adalah eigenface. Eigenface digunakan untuk melakukan ekstraksi ciri pada wajah. Berikut merupakan tahapan proses eigenface :

1. Citra training hasil konversi ke grayscale direduksi dimensi dari citra 2 D menjadi 1 D

sehingga membentuk vektor wajah yang merupakan vektor kolom.

2. Vektor-vektor tersebut disusun sedemikian rupa sehingga membentuk suatu matrik X dengan orde $n \times m$, dimana n adalah banyaknya jumlah piksel ($w \times h$) dan m adalah banyaknya citra training wajah.
3. Hasil ini selanjutnya akan digunakan untuk menghitung proses rata-rata setiap baris menggunakan rumus,

$$U_mean = \frac{1}{N} \sum_{i=1}^N x_i$$

4. Tahap selanjutnya dilakukan proses normalisasi pada data training. Normalisasi diperoleh dari rumus,

$$A_{(i,j)} = U_{(i,j)} - U_mean$$

5. Hasil dari normalisasi ini kemudian akan digunakan untuk membuat matriks kovarian (C) dimana

$$C = A^t * A.$$

6. Dari matriks kovarian ini diperoleh eigen vector (V) dan eigen value (D), dimana dalam bahasa matlab dituliskan dengan rumus,

$$C * V =$$

$$D * V$$

7. Eigen vector kemudian diurutkan berdasarkan eigen value dari yang terbesar sampai yang terkecil (V_Sort).
 8. Menghitung eigenface dimana, $eigenface = (A * V_Sort)^t$
- Setelah semua proses eigenface dilakukan, selanjutnya citra diambil cirinya melalui tahap eigenface PCA dengan menghitung *feature* PCA dari masing-masing

citra training yang dapat dihitung dengan rumus,

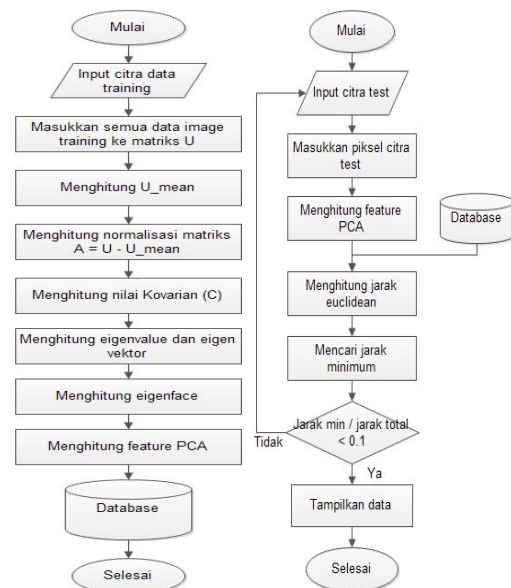
$$F_{(i,n)} = eigenface * A_{(i,n)}$$

Citra data training dan testing yang telah melewati tahap pengambilan ciri oleh PCA kemudian dihitung kesamaan jaraknya menggunakan perhitungan jarak Euclidean. Nilai minimum atau terkecil dari perhitungan jarak Euclidean *distance* merupakan nilai yang diambil sebagai hasil dari citra wajah yang dikenali.

4. Hasil Penelitian dan Pedmbahasan

4.1. Desain Proses

Sistem simulasi ekstraksi fitur untuk pengenalan wajah memiliki beberapa tahapan proses. Untuk mempermudah visualisasi penggunaan, berikut adalah *flowchart* dari sistem simulasi ekstraksi fitur.



Gambar 2. Desain proses

4.2. Pre-processing

Pre-processing ini bertujuan agar informasi yang terkandung pada citra data training yang diproses oleh

sistem dapat diolah dalam proses selanjutnya. Mula-mula citra training masukkan dikonversi menjadi citra grayscale.



Gambar 3. Hasil koversi RGB ke Grayscale

Setelah melalui tahapan konversi citra RGB menjadi citra grayscale, selanjutnya citra diproses melalui tahapan eigenface. Setelah nilai eigenface diperoleh, kini tugas dari PCA untuk mereduksi ciri yang masih terdapat pada citra data training. Hasil dari reduksi tersebut yaitu berupa pengurangan dimensi data, jadi dimensi data yang memiliki ciri yang dianggap tidak penting akan dihilangkan dan tidak akan dipakai dalam pemrosesan selanjutnya. Setelah nilai PCA dari semua citra data training diperoleh selanjutnya menghitung nilai PCA dari citra test.

4.3. Jarak Euclidean

Setelah nilai ciri PCA citra data training dan citra yang diuji diperoleh, langkah selanjutnya yaitu proses akhir. Pada tahap ini ciri citra data training dan ciri citra yang diuji dihitung kemiripannya dengan mencari minimum jarak kesamaannya. Perhitungan ini dilakukan dengan menggunakan jarak euclidean atau *euclidean distance*.

4.4. Hasil Pengujian

Berikut ini hasil dari beberapa pengujian yang telah dilakukan :

1. Pengujian pertama menggunakan 180 data training dimana setiap individu mempunyai 10 citra data training.

Tabel 1. Persentase keberhasilan setiap individu pada pengujian pertama

No	Nama	Persentase Keberhasilan
1	Adel	100%
2	Adit	80%
3	Agasa	60%
4	Agus	40%
5	Aldo	80%
6	Ari	100%
7	Bustan	60%
8	Dicky	60%
9	Duto	40%
10	Gunawan	60%
11	Kresna	100%
12	Dwiandi	100%
13	Nino	100%
14	Ommmy	100%
15	Retno	100%
16	Rio	60%
17	Rizky	100%
18	Tito	80%
Persentase rata-rata		78.89%

Dari tabel 1 didapat persentase keberhasilan rata-rata sebesar 78.89% yang didapat dari 90 citra uji. Terdapat 71 citra yang berhasil dikenali dengan baik oleh sistem, sementara 19 citra lainnya tidak dapat dikenali dengan baik yang

mungkin dikarenakan terjadinya kesalahan pada proses preprocessing maupun pada proses perhitungan jarak minimum antara citra yang diuji dengan citra data training.

2. Pengujian kedua menggunakan 108 data training dimana setiap individu mempunyai 6 citra data training.

Tabel 2. Persentase keberhasilan setiap individu pada pengujian kedua

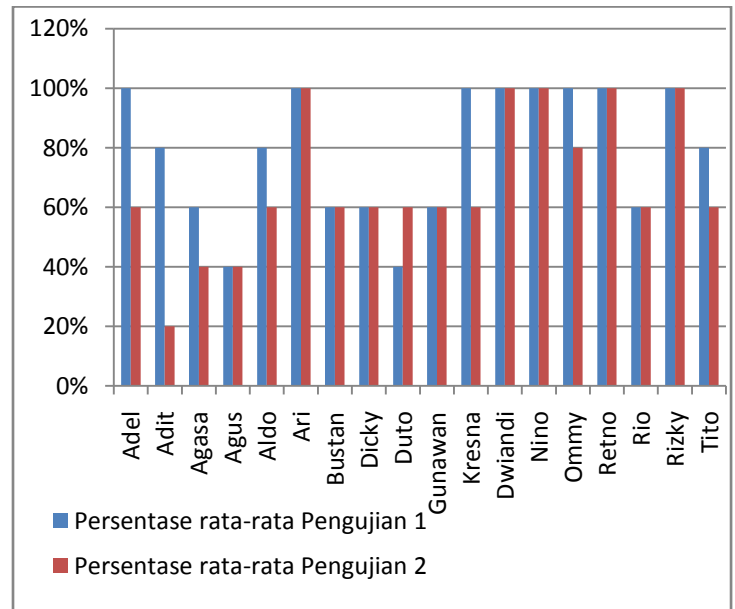
No	Nama	Persentase Keberhasilan
1	Adel	60%
2	Adit	20%
3	Agasa	40%
4	Agus	40%
5	Aldo	60%
6	Ari	100%
7	Bustan	60%
8	Dicky	60%
9	Duto	60%
10	Gunawan	60%
11	Kresna	60%
12	Dwiandi	100%
13	Nino	100%
14	Ommy	80%
15	Retno	100%
16	Rio	60%
17	Rizky	100%
18	Tito	60%

Persentase rata-rata	67.78%
-----------------------------	---------------

Dari tabel 2 didapat persentase keberhasilan rata-rata sebesar 67.78% yang didapat dari 90 citra uji. Terdapat 61 citra yang berhasil dikenali dengan baik oleh sistem, dan 29 citra lainnya tidak dapat dikenali dengan baik.

4.5. Analisa

Setelah semua citra telah diujikan, maka dapat dianalisa hasil dari proses ekstraksi fitur pada pengenalan wajah sistem. Berdasarkan hasil yang telah diperoleh, persentase akurasi keberhasilan sistem pada tabel 4.13. dan 4.14 didapatkan grafik perbandingan hasil pengujian yang terlihat pada gambar 4.13.



Gambar 4. Grafik persentase keberhasilan pengujian 1 dan 2

Berdasarkan hasil dari kedua pengujian, pada pengujian pertama memperlihatkan akurasi keberhasilan yang lebih tinggi dibandingkan dengan hasil pengujian kedua yaitu sebesar 78.89% dari 90 citra uji. Hal ini menunjukkan bahwa akurasi dipengaruhi oleh jumlah citra setiap individu pada data training, semakin banyak citra setiap orangnya dengan ekspresi yang berbeda pada data training maka akurasi yang diperoleh kemungkinan juga akan semakin besar dan begitu juga sebaliknya. Dari kedua pengujian yang telah dilakukan menunjukkan bahwa sistem yang telah dirancang berjalan cukup baik karena dari pengujian dapat dilihat kesalahan yang terjadi bukan hanya dikarenakan oleh kinerja proses sistem yang tidak baik. Namun terdapat beberapa faktor yang mungkin mempengaruhi kesalahan sistem. Faktor-faktor yang mempengaruhi kesalahan sistem dalam mengenali individu antara lain sebagai berikut :

1. Tempat, latar belakang dan tingkat kecerahan pengambilan citra
Dalam pengambilan citra masukan ini kemungkinan citra hasil pengambilan tidak memiliki tingkat kecerahan yang cukup baik karena diambil di tempat yang berbeda dengan intensitas cahaya yang berbeda pula sehingga mempengaruhi nilai yang didapat saat mengonversi citra dari RGB menjadi citra grayscale.

2. Kemiripan nilai eigenface tiap individu

Seperti halnya dengan mata manusia yang terkadang masih terjadi kesalahan dalam mengenali orang satu dengan yang lainnya. Begitu juga dengan sistem ini yang dalam pengambilan keputusan berdasarkan jarak minimum antara citra uji dengan citra data training yang terkadang setiap citra memiliki kemiripan dengan citra penghuni lain.

5. Kesimpulan dan Saran

5.1. Kesimpulan

Setelah dilakukan pengujian terhadap sistem simulasi ekstraksi fitur untuk pengenalan wajah, maka dapat diambil beberapa kesimpulan sebagai berikut :

1. Sistem simulasi ekstraksi fitur untuk pengenalan wajah dengan metode *Principal Component Analysis* (PCA) sebagai pengambilan ciri pada citra wajah dan *euclidean distance* untuk mencari minimum jarak antar citra data training dan citra data test menghasilkan nilai akurasi yang cukup baik sebesar 78.89%, namun membutuhkan data yang banyak sebagai data training.
2. Terdapat beberapa faktor yang mempengaruhi keberhasilan sistem yaitu tempat dan latar belakang pengambilan citra, tingkat kecerahan masukan citra serta kemiripan nilai eigenface tiap individu.

5.2. Saran

Berikut ini merupakan saran yang ditunjukkan untuk penelitian lebih lanjut :

1. Untuk menghasilkan tingkat akurasi yang lebih tinggi disarankan dalam pengambilan citra masukan diambil di satu lokasi yang sama dengan intensitas cahaya yang baik dan stabil.
2. Pada tahap preprocessing dapat ditambahkan fitur deteksi dan crop wajah sehingga waktu pemrosesan akan lebih singkat dan hasilnya akan lebih optimal.
3. Untuk mendapatkan hasil akurasi yang lebih optimal disarankan mengganti fitur ekstraksi dengan menggunakan fitur ekstraksi yang didasarkan pada bentuk dan penempatan atribut wajah, seperti mata, alis, hidung, bibir dan dagu serta hubungan antara atribut.

DAFTAR PUSTAKA

- [1] Suciati, U. 2007. Pengenalan Ekspresi Wajah dengan Ekstraksi Fitur Gabor Wavelet dan Metode Jaringan Syaraf Tiruan Backpropagation. Tugas Akhir Institut Teknologi Telkom.
- [2] Jannah, M. 2012. Sistem pengenalan wajah menggunakan jaringan syaraf tiruan algoritma backpropagation dengan metode eigenface principal component analysis. Tugas Akhir Universitas Pendidikan Indonesia.
- [3] Wikipedia. Ras manusia. Internet : http://id.wikipedia.org/wiki/Ras_manusia. (diakses pada tanggal 5 April 2013)
- [4] Nita Zelfia Dinianti. Sistem Pengenalan Citra Wajah. Internet : http://nita_zelfiafst09.web.unair.ac.id/artikel_detail-44283-Umum-Sistem%20Pengenalan%20Citra%20Wajah.html (diakses pada tanggal 29 Maret 2013)
- [5] Sholeh. A. 2013. Pengembangan Sistem Pengenalan Wajah 2D Dengan Implementasi Algoritma Eigenface dan Manhattan Distance. Tugas Akhir Universitas Pendidikan Indonesia
- [6] Aries, S.P. 2011. Deteksi Wajah Manusia dengan Metode Template Matching dan Ekstraksi Fitur Gabor. Tugas Akhir Institute Teknologi Telkom.
- [7] Lim, R. 2007. Pelacakan dan Pengenalan Wajah Menggunakan Webcam dan Metode Gabor Filter. Tugas Akhir Universitas Petra Christian.
- [8] Pearson, K. 1901. "*On Lines and Planes of Closest Fit to Systems of Points in Space*"(PDF). *Philosophical Magazine*, 2 (11): 559–572.
- [9] Abdi. H., & Williams, L.J. (2010). "*Principal component analysis*". *Wiley Interdisciplinary Reviews: Computational Statistics*, 2: 433–459.
- [10] Wikipedia. Principal Component Analysis. Internet : https://en.wikipedia.org/wiki/Principal_component_analysis (diakses pada tanggal 29 Maret 2013)
- [11] Pentland, A., Baback M. 1994. Face Recognition using View Based and Modular Eigenspaces. *Automatics System for the Identification and Inspection of Humans*, SPIE vol 2277, July.

- [12] Wikipedia. Analisis Komponen Utama. Internet : https://id.wikipedia.org/wiki/Analisis_komponen_utama (diakses pada tanggal 29 Maret 2013)
- [13] Johnson, Richard A & Wichern, Dean W. *Applied Multivariate Statistical Analysis* (New Jersey: Prentice-Hall International Inc, 1998). ISBN 0-13-080084-8.
- [14] Smith, L.I. *A Tutorial on Principal Component Analysis*. Internet : http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf (diakses pada tanggal 30 Maret 2013).
- [15] Wikipedia. Covariance. Internet : <http://en.wikipedia.org/wiki/Covariance> (diakses pada tanggal 30 Maret 2013).
- [16] Linear Algebrae. Eigenvalue and Eigenvektor. Internet : http://algebra.math.ust.hk/eigen/01_definition/lecture2.shtml (diakses pada tanggal 30 Maret 2013)
- [17] Diana, N.E. 2007. Sistem Temu Kembali Citra Wajah Berbasis Eigenface. Tugas akhir Universitas Indonesia.
- [18] Cahyadi, D. 2007. Ekstraksi dan kemiripan fitur mata pada sistem identifikasi buron. Tugas akhir Universitas Indonesia.
- [19] Agusthinta, D.R, Adang Suhedra, Hendra. Ekstraksi Fitur dan Segmentasi Wajah sebagai Semantik pada Sistem Pengenalan Wajah.
- [20] Sugiharto, Ardi. (2006). Pemrograman GUI dengan MATLAB. Yogyakarta : Andi.
- [21] Prasetyo, Eko. (2011). Pengolahan Citra Digital dan Aplikasinya menggunakan Matlab. Yogyakarta : Andi
- [22] PCA Based Face Recognition System. Internet : <http://www.mathworks.com/matlabcentral/fileexchange/17032>, (diakses pada tanggal 5 April 2013).
- [23] Teknik Reduksi Dimensi (Studi Kasus Principal Component Analysis). Internet : <http://www.ppi-tokodai.org/2012/04/29/teknik-reduksi-dimensi-studi-kasus-principal-component-analysis/> (diakses pada tanggal 5 April 2013).
- [24] Arif Muntasa. Ekstraksi Fitur. Internet : <http://arifmuntasa.wordpress.com/2012/05/12/ekstraksi-fitur/> (diakses pada tanggal 7 Mei 2013)
- [25] Sandra Agustyan P, Sigit Wasista, dkk. Sistem Absensi Mahasiswa secara Visual menggunakan webcam dengan dynamic timers warping. Instituit Teknologi Sepuluh Nopember Surabaya.
- [26] Tantri, P 2002. Opetimasi Pemrograman Kuadratis Dengan Metode Barisan Pendekatan Linear (Frank-Wolfe). Tugas Akhir Universitas Diponegoro.